

УДК 519.68

Е. С. Афонкина, М. А. Гришина, Н. Н. Ившина,
Г. А. Матвеев, В. А. Потемкин

Реализация параллельной версии программы расчёта лекарственных средств с использованием Т-Системы с открытой архитектурой (OpenTS)

Аннотация. В статье рассматривается реализация параллельной версии программы расчёта лекарственных средств (ПС ViS) с использованием Т-Системы с открытой архитектурой (OpenTS). Приведены результаты тестирования исходной (однопроцессорной) и параллельной (многопроцессорной) версий программы.

Ключевые слова и фразы: Т-Система, OpenTS, вычислительный кластер, язык параллельного программирования, лекарственное средство, биологическая активность соединений.

Введение

В статье используется следующий перечень сокращений и специальных терминов:

- OpenTS — Т-Система с открытой архитектурой;
- Т++ — язык параллельного программирования для Т-Системы, являющийся синтаксически и семантически гладким расширением C/C++ при помощи нескольких ключевых слов;
- Т-функция — функция, реализованная на языке Т++;
- ПС ViS — программная система ViS анализа биологической активности соединений и прогноза биологической активности новых потенциальных лекарственных средств;

Работа выполнена при поддержке научно-технической Программы Союзного государства «Разработка и использование программно-аппаратных средств Грид-технологий перспективных высокопроизводительных (суперкомпьютерных) вычислительных систем семейства «СКИФ» (шифр «СКИФ-ГРИД»), РФФИ (гранты РФФИ: № 07-03-96041, № 07-04-96053), Human Capital Foundation (London, Great Britain).

- рецептор — чувствительное нервное окончание или специализированная клетка, преобразующая воспринимаемое раздражение в нервные импульсы;
- лиганды — в комплексных соединениях — молекулы или ионы, непосредственно связанные с центральным атомом;
- конформация молекул — это геометрические формы, которые могут принимать молекулы органических соединений при вращении атомов или групп атомов;
- супрамолекулярная химия — это раздел химии, изучающий структуру и функции ассоциаций двух или более химических частиц, удерживаемых вместе межмолекулярными силами.

Системы параллельного программирования имеют важное прикладное значение. Это связано с быстрым развитием компьютерной техники, с переходом на многоядерные архитектуры, с развитием кластерных и распределённых систем. Для реализации параллельной версии программы расчёта лекарственных средств была выбрана система программирования OpenTS [1, 2], обладающая следующими важными характеристиками:

- совместимость со стандартными платформами и программными средствами;
- универсальность — система OpenTS применима в широком классе прикладных областей;
- масштабируемость в широком диапазоне числа процессоров;
- позволяет создавать эффективные приложения для компьютерных систем с высокой степенью параллелизма.

1. Описание алгоритма ViS

Программная система ViS позволяет определить строение комплексов «моделный рецептор-лиганд», установить взаимосвязь биологической активности с параметрами строения таких комплексов, выявить фармакофорные фрагменты молекулы. Последовательная версия программы реализована на языке Fortran.

Блок-схему алгоритма можно представить в виде следующих шагов:

- (1) ввод исходных данных: выборка соединений, координат атомов и величин биологических активностей соединений;
- (2) создание псевдоатомного окружения;

- (3) оптимизация ориентации молекул в псевдоатомном поле;
- (4) уточнение потенциалов поля (зарядов и радиусов псевдоатомов);
- (5) вывод результатов: молекулы, ориентированные в псевдоатомном рецепторе, энергии и силы взаимодействия в модельном комплексе «рецептор-лиганд», зависимость биологической активности от характеристик взаимодействия в системе «рецептор-лиганд».

На первом шаге строится псевдо-атомное окружение, комплементарное к наиболее активной молекуле выборки. В каждой точке поверхности молекулы можно определить потенциал кулоновских и ван-дер-ваальсовых взаимодействий и смоделировать комплементарное поле. Потенциал кулоновских взаимодействий в m -ой точке поверхности определяется по известной формуле:

$$\varphi_m^q = \sum_{i=1}^N \frac{q_i}{R_{im}} k,$$

где N — число атомов рассматриваемой структуры, q_i — заряд атома i , R_{im} — расстояние от точки m до атома i , k — коэффициент пересчёта в единицы СИ.

Для расчета потенциала ван-дер-ваальсовых взаимодействий, наводимых на точку m , по аналогии с кулоновским предложено использование следующей формулы:

$$\varphi_m^{VDW} = -2 \sum_{i=1}^N V_{im} \frac{2^3 r_i^3}{R_{im}^6},$$

где V_{im} — глубина минимума потенциальной энергии в соответствии с потенциалом Леннард-Джонса, r_i — ван-дер-ваальсовый радиус атома i . Расчет величин V_{im} и r_i может быть произведен в рамках модели MERA [3]. В данном уравнении для упрощения расчетов использован только член притяжения в ван-дер-ваальсовых взаимодействиях.

Комплементарное поле строится в виде псевдоатомного окружения молекулы. В точках поверхности строятся пробные сферы (псевдоатомы) с зарядами и радиусами, комплементарными полю в данной точке. Заряд и радиус пробной сферы вычисляются по формулам

$$q_m = -\frac{\varphi_m^q}{\sum_{i=1}^N \frac{k}{R_{im}}}, \quad r_m = \sqrt[3]{\frac{\varphi_m^{VDW}}{-2^3 \sum_{i=1}^N \frac{2V_{im}r_i^3}{R_{im}^6}}}.$$

На следующем шаге производится оптимизация ориентации второй молекулы в поле, комплементарном первой молекуле, с использованием комбинации симплексного и квази-ньютонского методов до достижения минимума совокупной вероятности контакта (P) её атомов со всеми точками модельного рецептора.

$$P = 1 - \prod_{m=1}^M (1 - p_m),$$

где

$$p_m = \exp\left(-\frac{E_m}{RT}\right),$$

$$E_m = \sum_{i=1}^N \left(\frac{kq_i q_m}{R_{im}} - 2V_{im} \frac{(r_m + r_i)^6}{R_{im}^6} + V_{im} \frac{(r_m + r_i)^{12}}{R_{im}^{12}} \right),$$

где N — число атомов во второй молекуле.

При этом вторая молекула может иметь размеры, значительно отличающиеся от размеров первой молекулы. Расположение пробных сфер на поверхности второй молекулы позволяет учесть конформационную подстройку рецептора под размер встраиваемой молекулы. В найденной ориентации второй молекулы уточняется комплементарное поле рецептора путём добавки потенциалов поля второй молекулы к имеющимся:

$$\varphi_m^q = \varphi_m^q + \varphi_m^q{}'; \quad \varphi_m^{VDW} = \varphi_m^{VDW} + \varphi_m^{VDW}{}'.$$

Потенциалы поля второй молекулы $\varphi_m^q{}'$ и $\varphi_m^{VDW}{}'$ вычисляются аналогично потенциалам первой молекулы. Аналогичная процедура осуществляется для третьей, четвертой и всех последующих молекул выборки. По окончании процедуры производится рассмотрение расположения первой, второй и т.д. молекул выборки в уточненном комплементарном поле. Процедура прекращается, если различие потенциала поля на новом и предыдущем шаге не превышает заданного предела. После этого строится линейная зависимость между величиной биологической активности соединения и параметрами взаимодействий псевдо-атомов модельного рецептора с молекулой. В качестве параметров взаимодействий могут быть использованы энергии взаимодействия и силы F_m :

$$F_m = \sum_{i=1}^N \frac{dE}{dR_m}.$$

Полученное уравнение позволяет выделить те псевдо-атомы модельного рецептора, взаимодействие с которыми определяет величину

биологической активности соединений. Данные псевдо-атомы моделируют активные центры рецептора. Фрагменты лекарственного средства, расположенные вблизи активных центров, определяют фармакофорную часть молекулы.

На основе предложенного алгоритма не сложно предложить новый подход к дизайну перспективных лекарственных средств. В рамках данного подхода производится комбинаторная замена атомов или фрагментов на другие атомы или фрагменты у наиболее активной молекулы с целью получения более активного соединения. Вновь полученное соединение помещается в модельный рецептор. На основе характеристик взаимодействия соединения с модельным рецептором вычисляется величина биологической активности. Если полученное значение находится на достаточно высоком уровне — перспективное соединение может быть синтезировано. Данный подход к дизайну перспективных лекарственных средств реализован в рамках алгоритма BiS.

Таким образом, алгоритм BiS позволяет определить строение комплексов «модельный рецептор-лиганд», установить взаимосвязь биологической активности с параметрами строения таких комплексов, выявить фармакофорные фрагменты молекулы.

2. Параллельная версия ПС BiS. Результаты тестирования

Для моделирования супрамолекулярных комплексов «рецептор-лекарственное средство» создан параллельный прототип ПС BiS с использованием Open^{TS}.

Алгоритм работы версии программы на T++ выглядит так: основная подпрограмма осуществляет поворот молекулы лекарственного средства в полярных координатах. При каждом изменении угла производится вызов T-функции, которая строит для молекулы модельный рецептор и вычисляет энергию взаимодействия в системе «модельный рецептор-лекарственное средство». Эта энергия записывается в выходные данные T-функции. Таким образом, осуществляется поиск наиболее выгодной (низкой по энергии) ориентации молекулы в модельном рецепторе.

Программа тестировалась на кластере «СКИФ МГУ» (Т60), имеющем следующую конфигурацию:

- операционная система: AltLinux НРС, x86_64;
- версия ядра ОС: 2.6.18;

ТАБЛИЦА 1. Результаты тестирования ПС ViS на кластере «СКИФ МГУ»

Количество процессоров	Время выполнения последовательной версии, сек	Время выполнения параллельной версии, сек	Ускорение
1	982,061	982,061	1,000
2		982,046	1,000
4		487,481	2,015
8		240,244	4,088
12		169,255	5,802
16		116,524	8,428
20		98,675	9,952
24		81,108	12,108
28		64,042	15,335
32		63,582	15,446
37		46,909	20,935
111		21,211	46,300

- число вычислительных узлов: 625;
- число процессоров: 1250;
- число ядер: 5000;
- модель процессора: Intel Xeon E5472 3.0 ГГц;
- число процессоров на узле: 2;
- число ядер на процессоре: 4;
- системная сеть: InfiniBand DDR.

Тестирование производилось на кластере с использованием от 1 до 111 процессоров. Результаты тестирования представлены в таблице 1.

Результаты расчётов параллельного прототипа ПС ViS совпадают с результатами последовательного аналога. Результаты моделирования супрамолекулярных комплексов «рецептор-лекарственное средство» в рамках алгоритма ViS сопоставлены с данными рентгеноструктурного анализа (РСА).

Расчёты проведены для агонистов $5HT_{1A}$ -рецептора (транквилизаторные средства) и ингибиторов дигидрофолатредуктазы (противоопухолевые и туберкулоостатические средства).

Выводы

Исследования показали следующее:

- параллельная реализация программы с помощью T-системы с открытой архитектурой (OpenTS) компактна и проста;
- потребовались лишь незначительные изменения исходного кода, чтобы сделать версию программы на языке T++;
- получено хорошее ускорение работы параллельной версии программы относительно её последовательного аналога.

Список литературы

- [1] Абрамов С. М., Адамович А. И., Инюхин А. В., Роганов В. А., Шевчук Ю. В., Шевчук Е. В. *T-система с открытой архитектурой* // Суперкомпьютерные системы и их применение SSA'2004: Труды Международной научной конференции, 26–28 октября 2004 г. Минск, ОИПИ НАН Беларуси. — Минск, 2004, с. 18–22. ↑(document)
- [2] Abramov S., Adamovich A., Inyukhin A., Roganov V., Shevchuk E., Shevchuk Yu., Vodomerov A. *OpenTS: An Outline of Dynamic Parallelization Approach* // РАСТ'2005, LNCS, 2005. ↑(document)
- [3] Потемкин В. А., Барташевич Е. В., Белик А. В. *Модель расчета атомных объемных характеристик в молекулярных системах* // Журнал физической химии. — т. 72, № 4, 1998, с. 650–656. ↑1

Челябинский Государственный Университет

Институт Органического Синтеза им. И.Я. Постовского УРО РАН

Исследовательский центр мультипроцессорных систем ИПС РАН

E. S. Afonkina, M. A. Grishina, N. N. Ivshina, G. A. Matveev, V. A. Potemkin. *Development and implementation of a parallel version of the algorithm «Biological Substrate Search» (BiS) using the T-System with the open architecture (OpenTS)* // Proceedings of Program Systems institute scientific conference “Program systems: Theory and applications”. — Pereslavl-Zalesskij, v. 1, 2009. — p. 217–223. — ISBN 978-5-901795-16-3 (in Russian).

ABSTRACT. The article considers development and implementation of a parallel version of the algorithm «Biological Substrate Search» (BiS) using the T-System with the open architecture (OpenTS). Tests results of a original (uniprocessor) and parallel (multiprocessor) versions of the algorithm showed.